

Expansion and Inclusiveness of Statistics

Simon Schwartzman

Invited paper prepared for session 83, "Statistics: an inclusive interpretation," of the 51st Session of the International Statistical Institute, 18-26 August 1997, Istanbul.

Summary

As the use of statistics expands in society, statisticians become more interested in the possibility of widening the scope of their discipline, to encompass and somehow bring this growing and unruly amalgam of users under control. Although this is, in appearance, an academic question - a rather abstract discussion on the scope and limits of a given field of enquiry - it has also very practical consequences for those of us who are engaged in the task of organizing and strengthening existing statistical institutions and offices, and providing society with better information and better ways of handling them. A brief outline of the factors leading to the expansion in the use of statistics is presented - its expansion to other disciplines, as a tool for public policy, and its growing availability due to the widespread use of computers. Then, there is a brief discussion of standard proposals to deal with the problem, based on articles and statements of leading personalities in the international Statistical community. In the conclusion, it is argued that statistics will probably not be redefined in a more inclusive way, but statisticians will have to play an increasing role of translators and brokers among different sectors in society that relate one way or another with the production and use of quantitative information. The pervasiveness of statistics is likely to continue to increase, and the statistician will have important roles to play, not only as translators and brokers, but also as creators of new meanings and providers of new instruments by which society can look and understand itself better.

It is not a coincidence that, as the use of statistics expands in society, statisticians become more interested in the possibility of widening the scope of their discipline, to encompass

and somehow bring this growing and unruly amalgam of users under control. Although this is, in appearance, an academic question - a rather abstract discussion on the scope and limits of a given field of enquiry - it has also very practical consequences for those of us who are engaged in the task of organizing and strengthening existing statistical institutions and offices, providing society with better information and better ways of handling them.

Expansion

The use of statistics is expanding in society in different ways. None of these tendencies is new, but, when brought together, they point to a potentially new situation. First, it is expanding to other scientific disciplines, so much so that it has been argued that statistical reasoning is the core "scientific method" which underlies all other fields of knowledge. Today, statistics is a central component of the education and daily works not only of demographers, sociologists and economists, but also of psychologists, epidemiologists, biomedical researchers, climatologists, and in many other disciplines. Each of these disciplines developed its own statistical traditions, giving preference to procedures and "rules of proof" which are not necessarily familiar or equally accepted by others. There is an obvious role and space, here, for broader statistical education of researchers in specialized fields, making them more aware of the limitations, strengths and possibilities of statistical reasoning and working procedures.

Secondly, statistical information leads to conclusions and decisions that can affect the lives of millions. This leads to efforts from the lay public to understand, question and influence the way data is collected and interpreted. In Brazil, population figures for municipalities affect their share of the country's tax income, and special credits they can get for education, health and poverty relief. Population figures, however, are not taken in isolation: they should be combined with data on regional income and the number of people living below the poverty line. Different estimations of the National Product can lead to different conclusions about a country's tax burden, and justify changes in tax. Unemployment statistics can affect the government's image, and lead to specific policies related to unemployment benefits, employment incentives or "flexibilization" of the job market. What is common in all these situations is that neither governments nor the

affected constituents are interested in the nuances, complexities and error margins that lie behind the single figures on population, product, income or poverty line

A third, more recent trend is the growing power of personal computers and the development of commercial computer software packages to perform statistical which were previously restricted to statistical offices endowed with large hardware, computer specialists and accomplished statisticians. Complex tabulations involving thousands of calculations and files with several hundred thousand megabytes can be done in a matter of minutes. Today, the user of such processing power has still to be able to choose among a wide array of strategies and to interpret the results. Progressively, however, artificial intelligence techniques will automate the procedures for selecting of statistical tools and interpreting the results. The consequence of this development is the growing demand for access to the micro-data produced by statistical offices, and the proliferation of summary data and results which would not necessarily receive the stamp of approval either from well-trained professional statisticians or from the statistical agencies responsible for the original data gathering.

There are other, less obvious aspects to this expansion. One has the clear impression that numbers are more important today in the press, and in shaping public opinion, than in the past. Rates related to economic growth, unemployment, inflation, income distribution, crime and deprivation produce headlines, affect the prestige of public figures, and can make and unmake ministers. Most of these rates did not exist before the current systems of official statistics were put in place. Some authors have speculated that this high profile of numbers is part of a more general trend to surround data with an aura of scientificity and objectivity, part of a broader trend to gain the trust and confidence of the population regarding scientists, researchers and their organizations (Porter, 1995). Statistical reasoning is starting to be introduced as elements of proof in legal proceedings, and statisticians are being called to perform roles similar to that of traditional forensic medicine in courts (Van-Matre and Clark, 1976; Gastwirth, 1988).

Standard Responses

There are two standard types of response to this situation, neither quite adequate. The first is to argue that statistics is a kind of "science of the sciences," and should therefore

incorporate and condition the ways other fields of knowledge are organized and behave. This is the well-known stand taken by Leslie Kish in his 1977 Presidential address (Kish, 1977) as well as by D.W. Marquardt a few years earlier in the same context (Marquardt, 1987). Closely associated is the notion that society should somehow restrict the use of statistical instruments and concepts to people without the proper training. The implicit goal is to control the use of statistical concepts and data the same way the medical profession controls the use of medical diagnostics and the consumption of prescription drugs.

The first ambition of statistics, for being the "primo inter pares" among sciences, can be dismissed without much trouble. Most fields of knowledge, from physics to sociology, share the same ambition. The trend, however, does not seem to be toward scientific convergence and unification, but toward increasing diversity and multiplicity of working paradigms. The efforts to control and limit the use of technical objects and concepts by laymen, best exemplified by the professional monopolies of medical doctors and lawyers, require extremely complex legal and institutional apparatus, and is probably on the wane even for medicine, challenged in the courts, through to the spreading of different forms of alternative medicine and deep changes in the traditional doctor-patient relationships, with consumer awareness replacing the usual relationships of trust and confidence. It is clear that, in this process, the medical profession is often on retreat, while the public is exposed to different forms of quackery; but it is not obvious that the current situation is worse than when the powers of the medical profession went unchallenged. It is possible that the professional monopoly of the legal profession will be the one to last longer, and not necessarily because of the scientific content of Law as an intellectual discipline.

Statisticians may aspire to the role of oracles, providing society with unquestioned interpretations and predictions of things to come, but such role is not easily granted. "In the absence of a culturally accepted mythology of deities to provide oracular status, statisticians can only gain status and recognition by providing value to their consumers. The statistician must assess what is wanted and needed by the consumer and find ways to provide both" (Boroto and Zahn, 1989). This stand creates an immediate reaction, which the authors are ready to acknowledge: "the idea of addressing what the consumer *wants*

(particularly if it conflicts with what the statistician thinks the consumer needs) could appear to be a form of prostitution." The answer, argue the authors, is not to provide the public with what it does not want, but to engage in active dialogue with the public - "the master statistician relies on dialogue"; an *effective* statistician is essentially a skillful translator, the consumer should never have the experience of being lost in a foreign land." A year later another article at *The American Statistician* took a similar position, arguing however for a much more active role for the statistician to perform. "As providers of information we can no longer complain that the users of our information do not know how to tell us what they need and, therefore, it is not our fault that if relevant, timely, easy-to-see and cost-effective information is not available for decision makers." The question, then, is not just to do what the consumer wants, but also to answer, to ourselves, "How can we govern, by our own actions, the future environments in which statisticians will work?" (Barabba, 1990).

Barabba's stand is a far cry from the technocratic arrogance of trying to shove one's preferred medicine down everyone else's throats, but it is nevertheless one step back from Boroto and Zahn's position. He is willing to make statistics more palatable, but he is the one who knows, for instance, that "we are entering a period in which the tolerable margin of error that both the governmental and private sectors will be allowed in the conduct of everyday affairs is greatly narrowed," and is willing to fulfill this need. Statisticians should be like good doctors: aware of the needs, concerns and limitations of their patients, but not to the point of letting them decide what afflicts them, and what medicine they should take.

Translators as Brokers

Is there a middle ground between the role models of the traditional doctor and that of the "shoe clerk" (Bross, 1974) who provides services at his costumers' demand?

I believe there is, and it has to with a subtle but important difference in the meaning of what "translation" means. When Barite and Zhan argue that the statistician should be a skillful translator, they mean that they should be able to explain to the laymen the meaning of statistical concepts, and translate in statistical terms the laymen's requests for

measurements and estimations. The assumption is that there is just one reality, expressed in different languages, the role of the translator being to make this unified reality explicit.

This is not, however, what translators do in real life. Languages are carriers of cultures, and cultures are never as close and exclusive as to exclude the possibility of translation of their meanings, but never so open and flat that everything can be transposed to other contexts without significant losses of meanings and content. This is true of people in different societies speaking different languages, and is also true of people occupying different positions in the same society - consumers, employees, statisticians, researchers, teachers, newspapermen, businessmen, and politicians. What does "unemployment" mean, for instance, for each of these people? Although they may share the same word, and understand it in ways that overlap, each thinks on unemployment in a way that is coherent with other aspects of their professional and social context. The role of the translator is not to make sure that everybody uses the term in the same way the statistician does, but to build bridges, and help each to understand the way others use the word.

Hopefully, this work of multiple translations can provide the basis for understandings and alliances that bring different sectors of society together, increase their mutual awareness, and bring benefits to all. Even if the meanings of "unemployment" are not the same to everybody, society can benefit from a stable, reliable and trustworthy source of unemployment statistics. It is the role of the statistician, as someone with a vested interest in such coalitions of interest, to go out of his way to understand the meanings and uses of statistical data by different groups in society, and build bridges between these needs and what the statistician can provide, thanks to his special skills and his knowledge of other, similar uses and needs. This requires a very active stand, different both from the passive response to the user's needs and from technocratic self-sufficiency. It requires to maintain one's knowledge, culture and approaches, and to develop the willingness to understand and accept the others' point of views.

The bottom line seems to be that statistics will probably not be redefined in a more inclusive way, from which it could recover the control over the uses society makes of its concepts and figures. But its pervasiveness will continue to increase, and the

statisticians will have important roles to play, not only as translators and brokers, but also as creators of new meanings and providers of new instruments by which society can look and understand itself better.

Bibliography

- Barabba, Vincent P., 1990 - "What does the future hold for statisticians?", *The American Statistician*, May, vol. 44, 2, 125-128.
- Boroto, Daniel R and Zahn, Douglas A., 1989 - "Promoting Statistics: on Becoming Valued and Utilized", *The American Statistician*, May, 43, 2, 71-72., p. 71.
- Bross, Irwin D. 1974 - "The Role of the Statistician: Scientist or Shoe Clerk", *American Statistician*; 28, 4, Nov, 126-127.
- Gastwirth, Joseph L., 1988 - *Statistical reasoning in law and public policy. Statistical modeling and decision science*. Boston: Academic Press,.
- Kish, Leslie, 1977 - "Chance, Statistics and Statisticians", *Journal of the American Statistical Association*, 73, 1-67.
- Marquardt, D.W. , 1987, "The Importance of Statisticians", *Journal of the American Statistician Association*, 82, 1-7.
- Porter, Theodore M., 1995 - *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton, N.J.: Princeton University Press.
- Van-Matre, Joseph G.; Clark, William N., 1976 - "The Statistician as Expert Witness", *American Statistician*; 30, 1, feb, 2-5 <